THE PERFORMANCE OF A SEQUENTIAL PROCEDURE FOR THE ESTIMATION OF THE MEAN

By Y. S. Chow¹ and K. F. Yu²

Columbia University and Yale University

Let $\{X_n, n \ge 1\}$ be a sequence of independent identically distributed random variables with mean μ and unknown variance σ^2 . We want to estimate μ by \bar{X}_n with a loss function of $\sigma^{2\delta-2}(\bar{X}_n-\mu)^2+\lambda n$, where $\delta>0$ and $\lambda\to 0+$. For $n_\lambda=o(\lambda^{-1/2})$ and $n_\lambda(\log\lambda)^{-1}\to -\infty$ as $\lambda\to 0+$, set $T=\inf\{n\ge n_\lambda: n^{-1}\sum_{i=1}^n (X_i-\bar{X}_n)+b_n\le \lambda^{1/\delta}a_n\}$. If $a_nn^{-2/\delta}\to 1$ and $0< b_n\to 0$ as $n\to\infty$, we prove that T is asymptotically risk efficient, that is, as $\lambda\to 0+$, $E[(2\lambda^{1/2}\sigma^\delta)^{-1}(\sigma^{2\delta-2}(\bar{X}_T-\mu)^2+\lambda T)]\to 1$. When the X_n 's are normal, the asymptotic risk efficiency of T was established by Starr. By introducing the delay factor n_λ , we are able to drop the condition of X_n 's being normal.

1. Introduction. Let X_1, X_2, \cdots be independent observations from some population with mean μ and variance σ^2 . With a sample (X_1, \dots, X_n) of size n, we want to estimate the unknown mean μ by \bar{X}_n with a loss structure

$$L_n = \sigma^{2\delta-2}(\bar{X}_n - \mu)^2 + \lambda n,$$

where $\bar{X}_n = n^{-1} \sum_{i=1}^n X_i$, $\delta > 0$ and $\lambda > 0$. The risk

$$R_n = EL_n = \frac{\sigma^{2\delta}}{n} + \lambda n$$

is minimized by taking a sample size n_0 , where

$$[\lambda^{-1/2}\sigma^{\delta}] \le n_0 \le [\lambda^{-1/2}\sigma^{\delta}] + 1, R_{n_0} = 2\lambda^{1/2}\sigma^{\delta} + 0(\lambda) \quad \text{as } \lambda \to 0+.$$

But if σ is unknown, there is no fixed sample size procedure that will attain the minimum risk. Robbins ([10], 1959, for $\delta = 1$) proposed to replace σ^2 by its estimator

$$V_n = (n-1)^{-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2$$

and to determine the sample size by

$$T' = \inf\{n \ge m : n \ge \lambda^{-1/2} V_n^{\delta/2}\}\$$

= \inf\{ n \ge m : V_n \le \lambda^{1/\delta} n^{2/\delta}\},

where $m \ge 2$ and to estimate μ by $\bar{X}_{T'}$. Let $R_{T'} = EL_{T'}$. The performance of T' is usually measured by

- (i) the risk efficiency: $R_{n_0}/R_{T'}$, and
- (ii) the regret: $R_{T'} R_{n_0}$.

Let the observations come from a normal population and $\delta = 1$. Using the loss structure

$$L_n^* = \lambda^{-1} ((\bar{X}_n - \mu)^2 + \lambda n).$$

Robbins [10] obtained some numerical and Monte Carlo results which suggested the boundedness of the regret, and Starr [11] established that, with $R_{n_0}^* = EL_{n_0}^*$ and $R_{T'}^* = EL_{T'}^*$, $R_{n_0}/R_{T'}$ = $R_{n_0}^*/R_{T'}^* \rightarrow 1$ as $\lambda \rightarrow 0$ (asymptotically risk efficient), if and only if $m \ge 3$. As $\lambda \rightarrow 0$, Starr

Received May 1979; revised September 1979.

¹ Research supported by the National Science Foundation under grant NSF MCS 76-09179.

² Research supported by Public Health under grant 5 RO1 GM 16895 at Columbia University. *AMS* 1970 subject classifications. Primary 62L12; secondary 60G40.

Key words and phrases. Uniform integrability, asymptotic risk efficiency.

and Woodroofe [12] found that the regret is bounded if and only if $m \ge 3$. Woodroofe [13] proved that if $m \ge 4$, then

$$ET' = \lambda^{-1/2}\sigma + \frac{1}{2}\sigma^{-2}\nu - \frac{3}{4} + o(1),$$

$$R_{T'}^* = 2\lambda^{-1/2}\sigma + \frac{1}{2}\nu + o(1).$$

where ν is a constant which can be computed. All these strongly indicate the good performance of the sequential procedure for the normal case. But is this procedure good in general? Let $P(X_1 = 1) = \mu = 1 - P(X_1 = 0), 0 < \mu < 1$, and $\delta = 1$. Then for $m \ge 2$

$$R_{T'} = E((\bar{X}_{T'} - \mu)^2 + \lambda T')$$

$$\geq \int_{\{X_1 = 1, \dots, X_m = 1\}} (\bar{X}_m - \mu)^2 dP$$

$$= (1 - \mu)^2 \mu^m > 0,$$

and $R_{n_0} \cong 2(\lambda \mu (1-\mu))^{1/2} \to 0$ as $\lambda \to 0$. Hence $\lim_{\lambda \to 0} R_{n_0}/R_{T'} = 0$ and T' is not asymptotically risk efficient.

To remedy the situation, Chow and Robbins [6] proposed the stopping rule (with $\delta = \beta = 1$)

$$T = \inf\{n \ge 2 : V_n + n^{-\beta} \le \lambda^{1/\delta} n^{2/\delta}\}, \qquad \beta > 0$$

by introducing the term $n^{-\beta}$. How good is the performance of the sequential procedure T? In this note, we shall prove the following theorem which shows that, no matter what the population is, T is asymptotically risk efficient as $\lambda \to 0$.

THEOREM. Let X, X_1, X_2, \cdots be independent identically distributed random variables with $EX = \mu$ and $Var\ X = \sigma^2 \in (0, \infty)$. For $\delta > 0$, let a_n and b_n be sequences of constants such that $a_n n^{-2/\delta} \to 1$ and $0 < b_n = o(1)$ as $n \to \infty$. For $\lambda > 0$ and $n_\lambda \ge 1$, define

$$T = T_{\lambda} = \inf\{n \ge n_{\lambda} : n^{-1} \sum_{k=1}^{n} (X_{k} - \bar{X}_{k})^{2} + b_{k} \le \lambda^{1/\delta} a_{k}\}.$$

Then,

- (i) If $n_{\lambda} = o(\lambda^{-1/2})$ as $\lambda \to 0$, then $\lim_{\lambda \to 0} \lambda^{1/2} T = \sigma^{\delta}$ a.s., and $\lim_{\lambda \to 0} E(\lambda^{1/2} T) = \sigma^{\delta}$.
- (ii) If $E|X|^{2p} < \infty$ for some p > 1 and $-K \log \lambda \le n_{\lambda} = o(\lambda^{-1/2})$ for some $K > K_{\alpha,p}$, then as $\lambda \to 0$,

$$\frac{R_T}{R_{n_0}} = E\left(\frac{\sigma^{2\delta-2}(\bar{X}_T - \mu)^2 + \lambda T}{2\lambda^{1/2}\sigma^{\delta}}\right) \to 1.$$

It should be pointed out that a key tool in proving this theorem under the mere assumption of a little bit more than the second moment on X is Lemma 5, whose proof will be given in detail in the next section.

2. Proof of the theorem.

LEMMA 1. Let $\{T \equiv T_{\lambda}, 1 \geq \lambda > 0\}$ be a family of random variables such that $P(T \geq 1) = 1$. If for some $0 < \gamma < 1$, p > 0, $\beta > 0$

(1)
$$P(\lambda^{\beta}T < \gamma) = o(\lambda^{\beta p}) \quad as \quad \lambda \to 0,$$

then $\{(\lambda^{\beta}T)^{-p}, 1 \ge \lambda > 0\}$ is uniformly integrable.

PROOF. Since $(\lambda^{\beta}T)^{-p}I_{[\lambda^{\beta}T\geq\gamma]} \leq \gamma^{-p}$ and by (1), $E(\lambda^{\beta}T)^{-p}I_{[\lambda^{\beta}T<\gamma]} \leq \lambda^{-\beta p}P(\lambda^{\beta}T<\gamma) = o(1)$, as $\lambda \to 0$, $\{(\lambda^{\beta}T)^{-p}, 1 \geq \lambda > 0\}$ is uniformly integrable.

Lemma 2. Let X, X_1 , X_2 , \cdots be independent identically distributed random variables with

 $EX = \mu \in (0, \infty)$ and $E(X^+)^p < \infty$ for some $p \ge 1$. Assume $\{a_n\}$ and $\{b_n\}$ are constants such that as $n \to \infty$, $b_n = o(1)$ and $n^{-\alpha}a_n \to 1$ for some $\alpha > 0$. Assume n_{λ} to be a positive integer-valued function of λ , α and p, with $n_{\lambda} = 0(\lambda^{-1/\alpha})$ as $\lambda \to 0$. For $\lambda > 0$, put

$$T \equiv T_{\lambda} = \inf\{n \ge n_{\lambda} : n^{-1}S_n + b_n \le \lambda a_n\},\$$

where $S_n = \sum_{j=1}^n X_j$. Then $\{(\lambda^{1/\alpha}T)^p, 1 \ge \lambda > 0\}$ is uniformly integrable.

PROOF. We can assume $\mu=1$. Define $\tau=\inf\{n\geq 1: S_n<2n\}$ and let $\tau_1,\,\tau_2,\,\cdots$ be the copies of τ , and $t_n=\tau_1+\cdots+\tau_n$. Then $S_{t_n}<2t_n$ and since $E(X^+)^p<\infty$, $E\tau^p<\infty$. (See [8]). By Doob's martingale theorem $\{(n^{-1}t_n)^p,\,n\geq 1\}$ is uniformly integrable. Choose a positive integer n_1 such that $a_n>\frac{1}{2}n^\alpha$ if $n\geq n_1$, and choose M such that $M\geq 2(2+c)$ and $M\geq \lambda n_\lambda^\alpha$, where $c\geq \sup_{n\geq 1}|b_n|$. Put

(2)
$$q = [(\lambda^{-1}M)^{1/\alpha}] + 1.$$

For $\lambda > 0$, and $q \ge n_1$, $t_q \ge q \ge n_\lambda$ and

$$t_q^{-1}S_{t_q} + b_{t_q} \le 2 + c \le \frac{2(2+c)}{\lambda} \frac{\lambda}{2} \le \frac{M}{\lambda} \frac{\lambda}{2} \le q^{\alpha} \frac{\lambda}{2} \le t_q^{\alpha} \frac{\lambda}{2} \le \lambda a_{t_q}.$$

Hence $T \le t_q$ a.s. By (2) the uniform integrability of $\{(n^{-1}t_n)^p, n \ge 1\}$ implies that of $\{(\lambda^{1/\alpha}t_q)^p, 1 \ge \lambda > 0\}$, and it follows that $\{(\lambda^{1/\alpha}T)^p, 1 \ge \lambda > 0\}$ is uniformly integrable.

LEMMA 3. Let X, X_1, X_2, \cdots be independent identically distributed random variables which are bounded from below with $EX = \mu \in (0, \infty)$. Let Y, Y_1, Y_2, \cdots be independent identically distributed bounded random variables with EY = 0. Assume $\{a_n\}$ and $\{b_n\}$ are constants such that as $n \to \infty$, $b_n = o(1)$ and $n^{-\alpha}a_n \to 1$ for some $\alpha > 0$. For $\lambda > 0$, let

$$T \equiv T_{\lambda} = \inf\{n \ge n_{\lambda} \colon n^{-1} \sum_{l=1}^{n} X_{l} - \bar{Y}_{n}^{2} + b_{n} \le \lambda a_{n}\},\$$

where $n_{\lambda} \ge -K \log \lambda$ for some K > 0. Then for p > 0, there exists $K_{\alpha,p} > 0$ such that if $K > K_{\alpha,p}$, $\{(\lambda^{1/\alpha}T)^{-p}, 1 \ge \lambda > 0\}$ is uniformly integrable.

The proof involves bounding the probability of $(\lambda^{1/2}T < \gamma)$ for $\gamma \in (0, 1)$ in the right order of λ . This can be done by a truncation argument and then applying the Kolmogorov's exponential bound. The details are omitted.

LEMMA 4. Let Z, Z_1 , Z_2 , \cdots be independent identically distributed random variables with variance $Var\ Z = \sigma^2$, which is positive and finite. Assume $\{a_n\}$ and $\{b_n\}$ are constants such that as $n \to \infty$, $b_n = o(1)$ and $n^{-\alpha}a_n \to 1$ for some $\alpha > 0$. For $\lambda > 0$, let

$$T = T_{\lambda} = \inf\{n \ge n_{\lambda} : n^{-1} \sum_{i=1}^{n} (Z_{i} - \bar{Z}_{n})^{2} + b_{n} \le \lambda a_{n}\},$$

where $n_{\lambda} \ge -K \log \lambda$, for some K > 0. Then for p > 0, there exists $K_{\alpha,p} > 0$ such that if $K > K_{\alpha,p}$, $\{(\lambda^{1/\alpha}T)^{-p}, 1 \ge \lambda > 0\}$ is uniformly integrable.

PROOF. For M > 0, put $Z'_n = (Z_n \wedge M) \vee (-M)$ for each $n \ge 1$, and choose M such that $Var Z'_n > 0$. For $\lambda > 0$, define

$$N \equiv N_{\lambda} = \inf\{n \ge n_{\lambda} : n^{-1} \sum_{i=1}^{n} (Z'_{i} - \bar{Z}'_{n})^{2} + b_{n} \le \lambda a_{n}\}.$$

Put $X_i = (Z_i' - EZ_i')^2$ and $Y_i = Z_i' - EZ_i'$. Then X_i and Y_i are bounded, and by Lemma 3, $\{(\lambda^{1/\alpha}N)^{-p}, 1 \ge \lambda > 0\}$ is uniformly integrable if $K > K_{\alpha,p}$. But by a result due to Chow and Studden [7],

$$\sum_{i=1}^{n} (Z_i' - \bar{Z}_n')^2 \le \sum_{i=1}^{n} (Z_i - \bar{Z}_n)^2.$$

Hence $N \le T$, and then $\{(\lambda^{1/\alpha}T)^{-p}, 1 \ge \lambda > 0\}$ is uniformly integrable.

LEMMA 5. Let Y_1, Y_2, \cdots be independent random variables with $EY_n = 0$ for each $n \ge 1$. Assume that for some $p \ge 2$, $\{|Y_n|^p, n \ge 1\}$ is uniformly integrable. Let \mathscr{F}_n be the σ -algebra generated by $\{Y_1, Y_2, \cdots, Y_n\}$ for each $n \ge 1$, $\mathscr{F}_0 = \{\phi, \Omega\}$, and let $\{M(b), b \in B\}$ be \mathscr{F}_n -stopping times with $B \subset (0, \infty)$ such that $\{(b^{-1}M(b))^{p/2}, b \in B\}$ is uniformly integrable. Let $W_n = \sum_{i=1}^n Y_i$. Then $\{|b^{-1/2}W_{M(b)}|^p, b \in B\}$ is uniformly integrable.

PROOF. For each $b \in B$, let $M' \equiv M'(b) = M(b) \land N$, where N = [Kb] for $K \ge 1$. Put $M_n = M' \land n$. Since $\{|Y_n|^p, n \ge 1\}$ is uniformly integrable, for any $\delta > 0$, there exists a positive constant K_1 such that

$$\sup_{n\geq 1} E | Y_n I_{\{|Y_n|\geq K_1\}} - E Y_n I_{\{|Y_n|\geq K_1\}} |^p < \delta.$$

Put $X_n = Y_n I_{\{|Y_n| \ge K_1\}} - E Y_n I_{\{|Y_n| \ge K_1\}}$ and $Z_n = Y_n - X_n$. Then

$$\sum_{i=1}^{M_n} X_i = \sum_{i=1}^n X_i I_{[M' \ge i]},$$

and

$$\sum_{i=1}^{M_n} Z_i = \sum_{i=1}^n Z_i I_{[M' \ge i]}$$

are martingales. By a result of Burkholder, Davis and Gundy [2], for some constant A > 0,

(3)
$$E \left| \sum_{i=1}^{M'} X_{i} \right|^{p} \leq AE(\sum_{i=1}^{\infty} E(X_{i}^{2} I_{[M' \geq i]} / \mathscr{F}_{i-1}))^{p/2} + AE(\sup_{i \geq 1} \left| X_{i} \right| I_{[M' \geq i]})^{p}$$

$$\leq AE(\sum_{i=1}^{\infty} I_{[M' \geq i]} EX_{i}^{2})^{p/2} + AE(\sum_{i=1}^{\infty} E \left| X_{i} \right|^{p} I_{[M' \geq i]})$$

$$\leq A\delta E(M')^{p/2} + A\delta E(M')$$

$$\leq 2A\delta E(M')^{p/2}.$$

Similarly

$$E |\sum_{i=1}^{M'} Z_i|^{p+1} \le 2^{p+2} A K_1^{p+1} E(M')^{(p+1)/2}$$
.

Therefore

$$\sup_{b \in B} E |b^{-1/2} \sum_{i=1}^{M'} X_i|^p \le 2A\delta \sup_{b \in B} E(b^{-1}M')^{p/2}$$

which can be made arbitrarily small; and

$$\sup_{b \in B} E \left| b^{-1/2} \sum_{i=1}^{M'} Z_i \right|^{p+1} \le 2^{p+2} A K_1^{p+1} K^{(p+1)/2} < \infty.$$

Hence

(4)
$$\{|b^{-1/2}\sum_{i=1}^{M'}Y_i|^p, b \in B\}$$

is uniformly integrable. Now

(5)
$$W_{M(b)} = \sum_{i=1}^{\infty} Y_i I_{[M(b) \ge i]}$$
$$= \sum_{i=1}^{N} Y_i I_{[M(b) \ge i]} + \sum_{i=N+1}^{\infty} Y_i I_{[M(b) \ge i]}.$$

As in the derivation in (3),

$$E \left| \sum_{i=N+1}^{\infty} Y_i I_{[M(b) \ge i]} \right|^p \le A E \left(\sum_{i=N+1}^{\infty} I_{[M(b) \ge i]} E Y_i^2 \right)^{p/2} + A E \left(\sup_{i \ge N+1} \left| Y_i \right|^p I_{[M(b) \ge i]} \right)$$

$$\le 2A \sup_{n \ge 1} E \left| Y_n \right|^p \int_{M(b) \ge N} (M(b))^{p/2} dP$$

$$= o(b^{p/2})$$

uniformly in b as $K \to \infty$. This, together with (4) and (5), completes the proof that

$$\{|b^{-1/2}W_{M(b)}|^p, b \in B\}$$

is uniformly integrable.

PROOF OF THE THEOREM. Obviously $T < \infty$ a.s., and as $\lambda \to 0$, $T \to \infty$ a.s. By the strong law of large numbers,

$$T^{-1} \sum_{i=1}^{T} (X_i - \bar{X}_T)^2 + b_T \rightarrow \sigma^2$$
 a.s.

Therefore $\lambda^{1/\delta} T^{2/\delta} \to \sigma^2$ a.s., and hence $\lambda^{1/2} T \to \sigma^{\delta}$ a.s. Define

$$\tau \equiv \tau_{\lambda} = \inf\{n \geq n_{\lambda} : n^{-1} \sum_{i=1}^{n} (X_i - \mu)^2 + b_n \leq \lambda^{1/\delta} a_n\}.$$

Let $\lambda' = \lambda^{1/\delta}$. Then $n_{\lambda} = O(\lambda'^{-\delta/2})$ and by Lemma 2 $\{\lambda^{1/2}\tau, 1 > \lambda > 0\}$ is uniformly integrable. Since $T \le \tau$, $\{\lambda^{1/2}T, 1 \ge \lambda > 0\}$ is uniformly integrable; consequently $E(\lambda^{1/2}T) \to \sigma^{\delta}$ as $\lambda \to 0$. This proves (i). For (ii), since $E|X|^{2p} < \infty$ and $n_{\lambda} = O(\lambda^{-1/2})$, by Lemma 2, $\{(\lambda^{1/2}\tau)^p, 1 \ge \lambda > 0\}$ and hence $\{(\lambda^{1/2}T)^p, 1 \ge \lambda > 0\}$ are uniformly integrable. By Hölder's inequality, Lemmas 4 and 5

$$E\left| (\lambda^{1/4} \sum_{i=1}^{T} (X_i - \mu))^2 \left(\frac{1}{\lambda T^2} - \sigma^{-2\delta} \right) \right| \le E^{1/p} |\lambda^{1/4} \sum_{i=1}^{T} (X_i - \mu)|^{2p} E^{1/q} |(\lambda T^2)^{-1} - \sigma^{-2\delta}|^q$$

$$= O(1) \cdot o(1) = o(1),$$

where p + q = pq. Hence by Lemma 5,

$$\lambda^{-1/2} (\bar{X}_T - \mu)^2 = (\lambda^{1/4} \sum_{i=1}^T (X_i - \mu))^2 \left(\frac{1}{\lambda T^2} - \sigma^{-2\delta} \right) + \sigma^{-2\delta} (\lambda^{1/4} \sum_{i=1}^T (X_i - \mu))^2$$

is uniformly integrable; consequently, together with Anscombe's theorem [1],

$$\lambda^{-1/4} \sigma^{\delta/2-1} (\bar{X}_T - \mu) \stackrel{\mathscr{L}}{\to} N(0, 1), \quad \text{as} \quad \lambda \to 0,$$

and (i), we have

$$\lim_{\lambda \to 0} \frac{R_T}{R_{n_0}} = \lim_{\lambda \to 0} E \left[\frac{\sigma^{2\delta - 2} (\bar{X}_T - \mu)^2 + \lambda T}{2\lambda^{1/2} \sigma^{\delta}} \right] = 1.$$

REMARKS. It should be noted that if the type of distribution for the population is known, then by direct computation, n_{λ} can be replaced by a fixed positive integer so that the conditions of Lemma 1 hold. In particular, the case when the X_n 's are normally distributed is certainly subsumed under our general results for asymptotic risk efficiency.

REFERENCES

- Anscombe, F. J. (1952). Large sample theory of sequential estimation. Proc. Cambridge Philos. Soc. 48 600-607.
- [2] BURKHOLDER, D. L., DAVIS, B. J. and GUNDY, R. F. (1972). Integral inequalities for convex functions of operators on martingales. *Proc. Sixth Berkeley Symp. Math. Stat. Prob.* 2 223-240.
- [3] CABILIO, P. and ROBBINS, H. E. (1975). Sequential estimation of p with squared relative error loss. Proc. Nat. Acad. Sci. USA 72 191-193.
- [4] CHOW, Y. S. (1971). On the L_p -convergence for $S_n/n^{1/p}$, 0 . Ann. Math. Statist. 42 187–190.
- [5] CHOW, Y. S., HSIUNG, C. A., and LAI, T. L. (1979). Extended renewal theory and mean convergence in Anscombe's theorem. Ann. Probability 7 304-318.
- [6] CHOW, Y. S., and ROBBINS, H. E. (1965). On the asymptotic theory of fixed width sequential confidence intervals for the mean. Ann. Math. Statist. 36 457-462.
- [7] CHOW, Y. S., and STUDDEN, W. J. (1969). Monotonicity of the variance under truncation and variation of Jensen's inequality. Ann. Math. Statist. 40 1106-1108.
- [8] Gut, A. (1974). On the moments and limit distributions of some first passage times. Ann. Probability 2 277-308.
- [9] Loève, M. (1977). Probability Theory I, 4th ed. Springer, New York, Heidelberg, Berlin.
- [10] ROBBINS, H. E. (1959). Sequential estimation of the mean of a normal population. Probability and Statistics (Harald Cramér Volume). Almquist and Wiksell, Upsala.

- [11] STARR, N. (1966). On the asymptotic efficiency of a sequential procedure for estimating the mean. Ann. Math. Statist. 37 1173-1185.
- [12] STARR, N., and WOODROOFE, M. (1969). Remarks on sequential point estimation. Proc. Nat. Acad. Sci. USA 63 285-288.
- [13] WOODROOFE, M. (1977). Second order approximation for sequential point and interval estimation.

 Ann. Statist. 5 984-995.

DEPARTMENT OF MATHEMATICAL STATISTICS COLUMBIA UNIVERSITY
NEW YORK, NEW YORK 10027

DEPARTMENT OF STATISTICS
YALE UNIVERSITY
BOX 2179 YALE STATION
NEW HAVEN, CONNECTICUT 06520